Analysis of Symmetric algorithm for XML document security

Nithin N
Department of Computer Science Engineering
SDMIT, Ujire, Karnataka, India

Harshitha.K.S.
Department of Computer Science Engineering
SDMIT, Ujire, Karnataka, India

Divyashree K
Department of Computer Science Engineering
SDMIT, Ujire, Karnataka, India

Shruti.N.Nayak
Department of Computer Science Engineering
SDMIT, Ujire, Karnataka, India

Abstract - The encryption machine takes the key Value and the input file (XML file) and generate the encrypted text using Symmetric algorithm (Caesar cipher and Vigenere cipher).The decryption Machine takes the encrypted file and key Value to generate the original XML file. It also checks whether the input (key value) is valid and block the user if invalid input is provided (key value). These algorithms are compared with respect to different granularity levels of XML files with various file sizes, time obtained for encryption and decryption.

Keywords – Caesar Cipher, Vigenere Cipher, XML, XML Granularity

I. INTRODUCTION

Cryptography plays a vital role in communication security. It is becoming increasingly important as a basic building block for computer security [1]. Cryptography is a practice and study of masking information. The word "Cryptography" is inferred from the Greek word 'kryptos', meaning hidden [2, 3]. An original message is known as plaintext, while the encoded message is called the cipher text. The process of converting plaintext (XML file) to cipher text (string or decimals) is known as encryption, restoring the plaintext from the cipher text is called decryption.

Extensible Mark-up Language (XML) has emerged rapidly as a new approach to delivering structured data over the web [4, 5].It is a markup language that is developed by the World Wide Web (WWW) Consortium to overcome the HTML (Hypertext Markup Language) limitations. In general, XML is a language for describing data on the web.XML document contains tags called as Mark-ups, which describe the content of the document. XML is extensible so it can be used to create many different applications.

Security is needed in XML document transaction for virtually all businesses, most government agencies, and many individuals now have Web sites. The number of individuals and companies with Internet access is expanding rapidly and all of these have graphical Web browsers. As a result, businesses are enthusiastic about setting up facilities on the Web for electronic commerce. But the reality is that the Internet and the Web are extremely vulnerable to compromises of various sorts. As businesses wake up to this reality, the demand for secure Web services grows.

Cryptographic algorithms are mainly classified into two types: Symmetric cryptography and Asymmetric cryptography [6]. Symmetric encryption is a form of cryptosystem in which encryption and decryption are performed using the same key. It is also known as conventional encryption. Asymmetric encryption is a form of cryptosystem in

which encryption and decryption are performed using the different keys- one a public key and one a private key. It is also known as public-key encryption.

This paper involves comparison of efficiencies of two symmetric encryption algorithms namely Caesar cipher and Vigenere cipher. In Caesar cipher, each letter of the alphabet is shifted along some number of places [7]. The Vigenere cipher is a method of encrypting alphabetic text by using a series of different Caesar ciphers based on the letters of the keyword. It is a simple form of polyalphabetic substitution [8].

The rest of the paper is organized as follows. Related work description in section II. The section III describes methodology of Caesar cipher and Vigenere cipher algorithms, Experimental setup is mentioned in section IV and results analysis of the algorithms are in section V. Section VI concludes the paper.

II. Related Work

Now a day's XML has been becoming popular for document representation and exchange over the Web. Since Security mechanisms for the protection of XML document sources and their distribution are essential [9]. There are some of the related work based on the security of the XML document and their distribution.

In paper [10], authors propose a cryptosystem (encrypting/decryption) for XML data using Vigenere cipher algorithm and EL Gamal cryptosystem. Such a system is designed to achieve some of security aspects such as confidentiality, authentication, integrity, and non-repudiation. We used XML data as an experimental work. Since, we have used Vigenere cipher which is not monoalphabetic, and then the number of possible keywords of length m in a Vigenere Cipher is 26m, so even for relatively small values of m, an exhaustive key search would require a long time.

Data security is the overriding concern in DAS, and encryption is a natural solution. However, queries over encrypted databases are usually inefficient due to a heap of time on the encryption and decryption. Authors present an encryption scheme and XQuery translation model of XML database in [11]. The major work and contribution of this paper are as below: 1) Splitting method is proposed in the encryption scheme, in which more flexible encryption granularity is obtained. 2) The Authors put forward the encryption strategy that, by encrypting two or more fragments together, is efficient to resist various kinds of database attack owing to the changed cipher text distribution and data size by splitting. 3) An XQuery translation model is introduced, by converting the query of XML data into that of relational data, combined with XML schemas compression and hash technology, which perform the XQuery language efficiently. Our experimental evaluation shows that our XML database encryption and query model achieves both excellent query efficiency and robust security.

In [12] proposes a cryptosystem (encrypting/decryption) for XML data using RSA (Rivest, Shamir, and Adleman) with some form of shift ciphering scheme. Such a system is designed to achieve some of security aspects such as confidentiality, authentication, and integrity, and non-repudiation. Author used XML data as an experimental work. The implementation is done using VB.NET. Since, Author have used RSA with some padding scheme; it is extremely difficult to factor large numbers. The property of shift ciphering scheme increases the cost of crypto analysis. The results are very much satisfactory for securing XML data. Author found the estimation required time to break our generated keys is 3128 years, which is sufficient against any brute-force attacks.

Jammalamadaka et. al. in [13] proposed a technique to query encrypted XML documents. Such a problem predominantly occurs in "database as a service" (DAS) architectures, where a client may outsource data to a service provider that provides data management services. Security is of paramount concern, as the service provider itself may be untrusted. Encryption offers a natural solution to preserve the confidentiality of the client's data. The challenge now is to execute queries over the encrypted data, without decrypting them at the server side. In this paper authors developed: (1) primitives using which a client can specify the sensitive parts of the XML documents; (2) mechanisms to map the XML documents to encrypted representation that hides sensitive portions of the documents; and (3) techniques to run SPJ (selection-projection-join) queries over encrypted XML documents. A strategy, where indices/ancillary information is maintained along with the encrypted XML documents is exploited, which helps in pruning the search space during query processing

In [14], presents a scheme for securing XML documents and their distribution. Our scheme has several advantages over Author-X such as: (a) one user needs only one private key; (b) even when the user leaves or a credential is changed, all the other users will be unaffected; (c) there is no need to establish a secure channel for key distribution; and (d) there is no need for checking the XML documents for access control policies applied. These make the security model more efficient and robust as well as simplifying the programming and the generation of the encrypted document base.

XML is used to exchange messages across different networks due to its flexibility and adaptability. Due to the wide usage of XML documents over internet there arises a need to protect these documents to provide data

confidentiality and privacy. In [15] author proposing a new public key cryptographic algorithm called XML Batch Multi-Prime(XBMRSA) to encrypt sections of XML document based on Multi-Prime RSA technique. A detailed comparative study of standard RSA and XBMRSA is carried out by considering parameters like required to encrypt and decrypt the XML file and size of encrypted file.

III. METHEDOLOGY

A. Caesar Cipher–

Caesar cipher is also known as shift cipher or Caesar shift [16]. It is the one of the simplest and most widely known encryption technique. One of the earliest know example for substitution cipher. It is said to have been used by Julius Caesar to communicate with his Army. In Caesar cipher each character of the plaintext message is replaced by a character n position down in the alphabet. It is one type of Symmetric key Cryptography. In this case the relationship between a character in the plaintext and character in the cipher text is always one-to-one. Figure 1 and Figure 2 shows the encryption and decryption steps involved in Caesar Cipher.

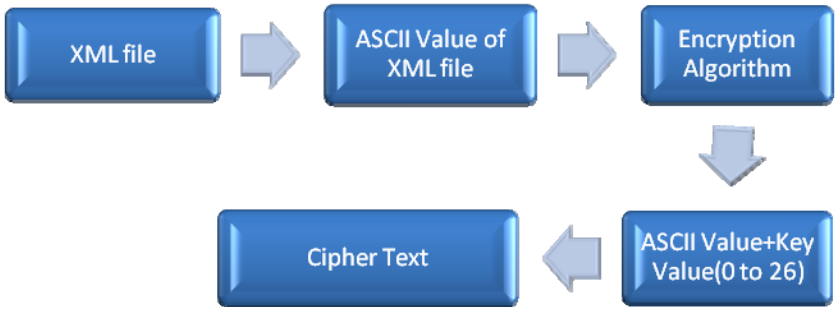

Figure 1. Caesar Cipher encryption steps

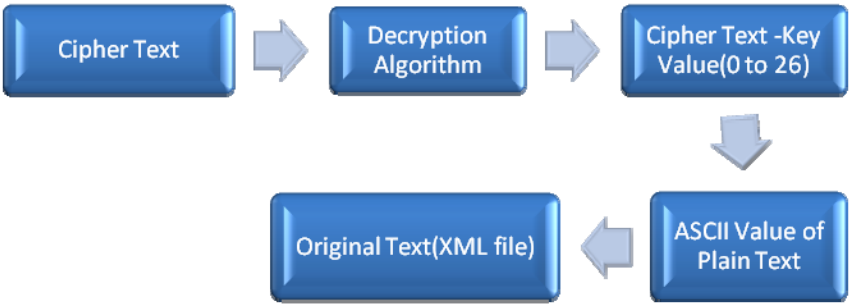

Figure 2. Caesar Cipher decryption steps

B. Vigenere Cipher–

The Vigenère cipher was invented by a Frenchman, Blaise de Vigenère in the 16th century [17]. Vigenere cipher is the simplest polyalphabetic substitution cipher uses two or more cipher alphabets to encrypt the data. It is the method of encrypting the series of text using multiple Caesar cipher.

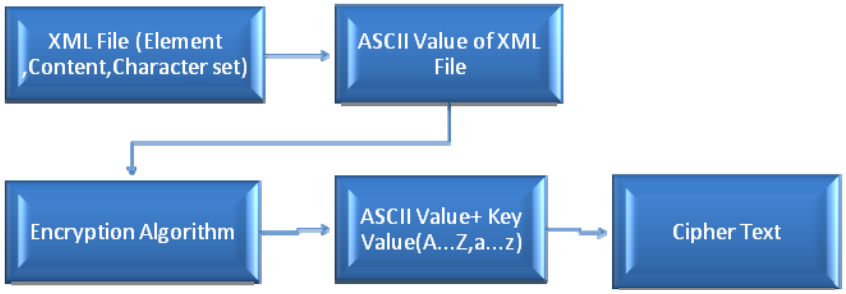

Figure 3. Vigenere Cipher for Encryption steps

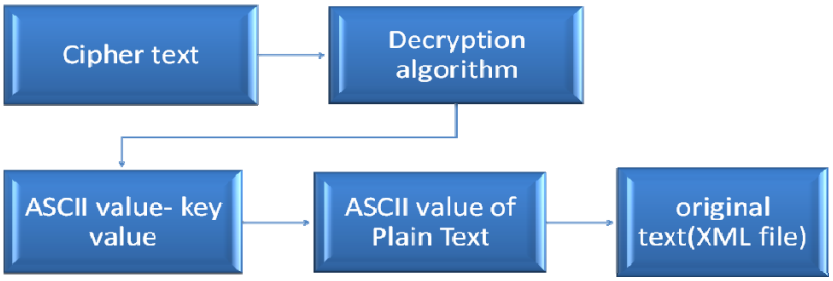

Figure 4. Vigener cipher for Decryption steps

IV. Experimental Setup

In this paper the results are obtained by simulating the entire algorithm in C# .Net 4.0 .Tool used for execution is Microsoft Visual Studio 2010. Operating system is Windows 7 and 1GB RAM and size of the XML file varies from 50KB to 500KB.

V. Result Analysis

Figure 5 and Figure 6 shows the encryption and decryption time (Total Time) for elements in the XML file for different key values for Vigenere Cipher and Caesar Cipher. As size of file increases, time taken for encryption and decryption time also increases. Since content and character set changes from file to file there may be variation in encryption and decryption time.

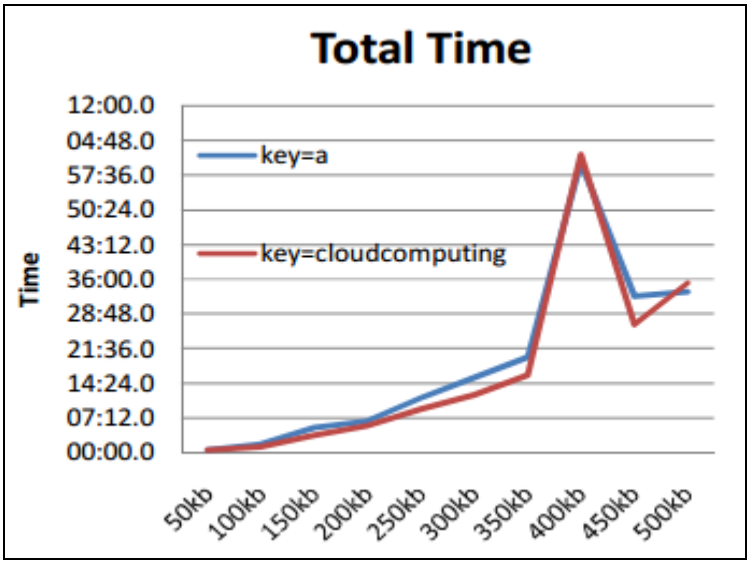

Figure 5. Total Time Taken by Vigenere cipher for Element

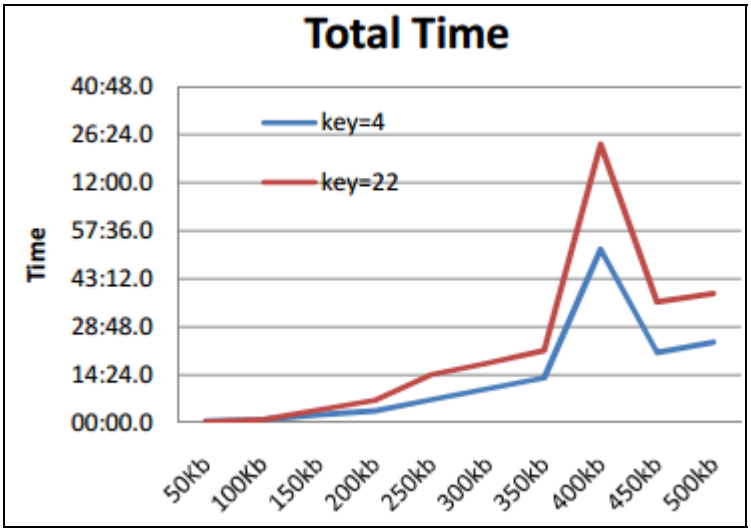

Figure 6. Total Time Taken by Caesar cipher for Element

Figure 7 shows the total time taken by key value 'ab' and 'cloudcomputing' for Content in XML File. The Figure 8 indicates the total time taken (encryption + decryption) by Caesar cipher for key value 22 and 4 respectively.

The Figure 9 and Figure 10 shows, the total time taken (encryption +decryption) for the character set in the XML document for some file the total time taken for different key value for Vigenere Cipher and Caesar Cipher respectively. Figure 13 shows that from browser to browser the encryption and decryption time varies. For example encryption time taken for file size 200kb Internet Explorer is 04:22.5, Google chrome is 04:43.5 and Mozilla Firefox is 03:47.0.

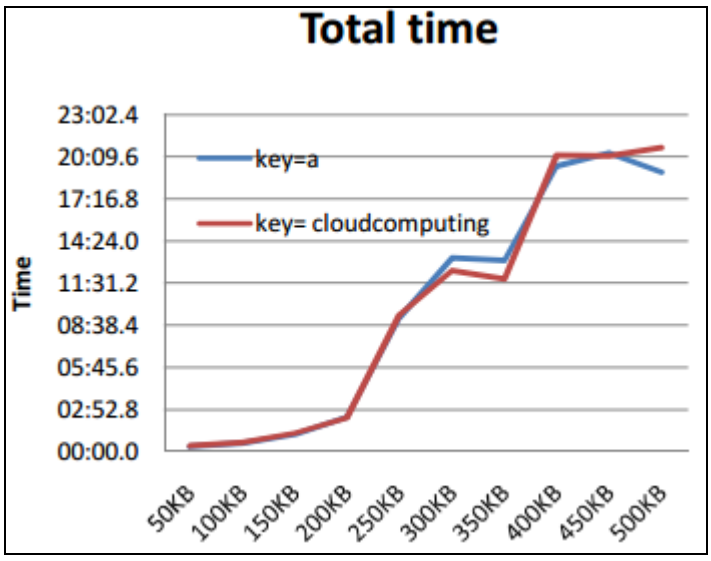

Figure 7. Total Time Taken by Vigenere cipher for Content

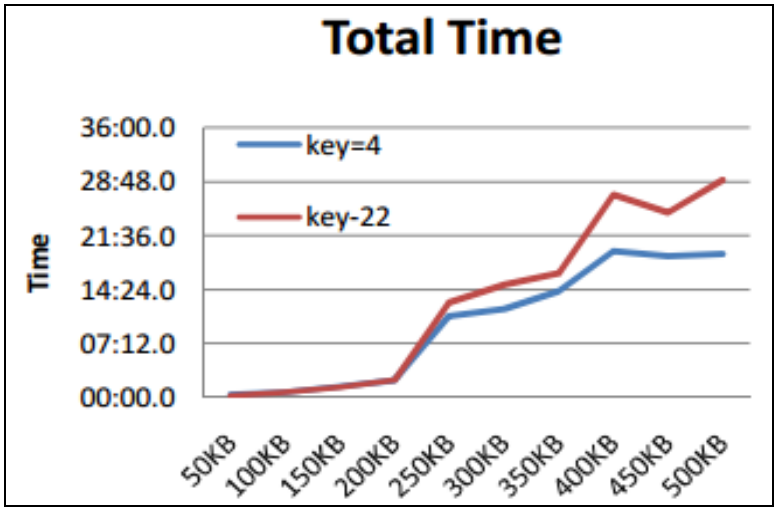

Figure 8. Total Time Taken by Caesar cipher for Content

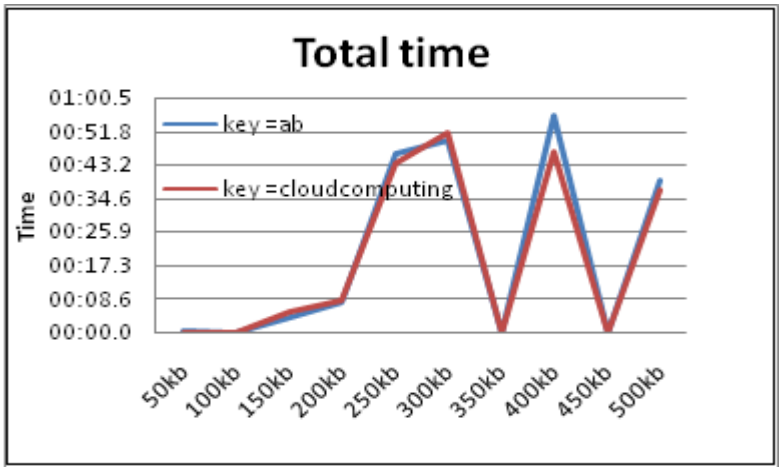

Figure 9. Total Time Taken byVigenere cipher for Character Set

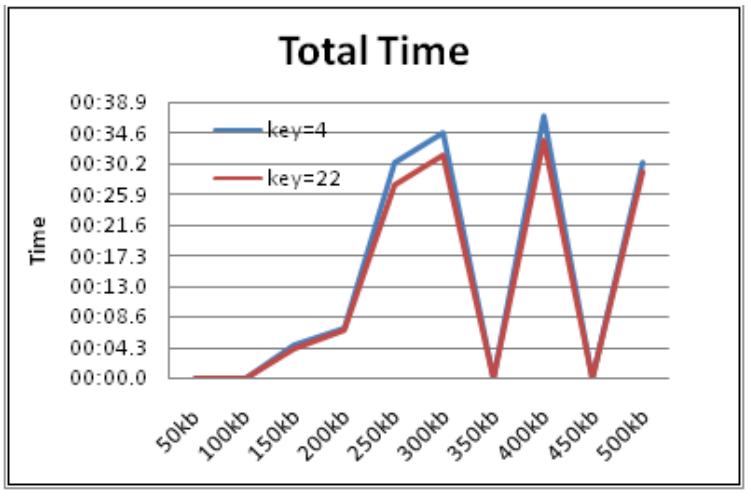

Figure 10. Total Time Taken by Caesar cipher for Character Set

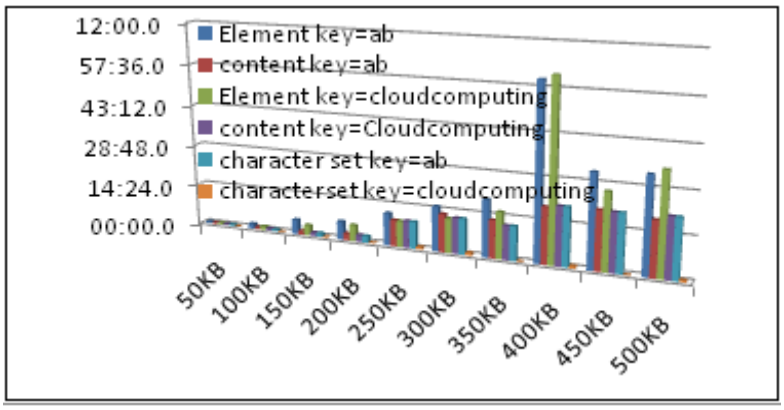

Figure 11. Total Time Taken by Vigenere cipher for XML file

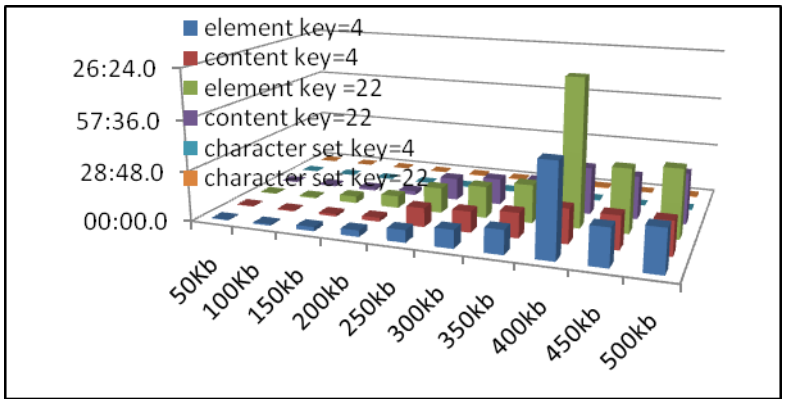

Figure 12. Total Time Taken by Caesar cipher for XML file

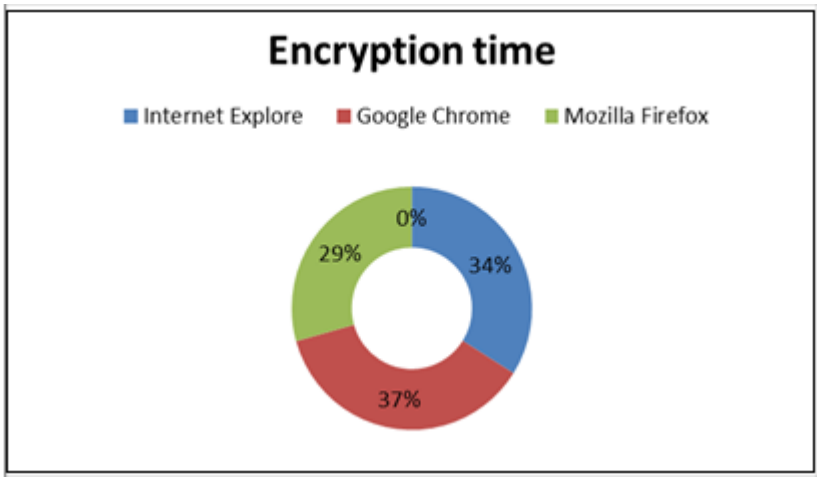

Figure 13. Vigener cipher for Decryption steps

IV.CONCLUSION

The result analysis shows that according to encryption and decryption time taken for Element, Content and Character Set for file size (50kb to 500kb). As the file size increases the encryption and decryption time increases. The time taken by Caesar Cipher to Encryption and Decryption is less when compare to Vigenere Cipher .The Vigenere Cipher cannot be easily breakable due to their polyalphabetic property. Vigenere cipher is more complex when compared to Caesar cipher.

V. REFERENCE

[1] Janailin Warjri, Dr.E.Gerorge Dharma and Prakash Raj, "Analysis Of Symmetric Key Algorithms", *International journal of societal applications of computer science*, Vol. 2, Issue. 9, pp. 454-457, September 2013.

[2] William Stalling, "Cryptography and Network Security", Published by Pearson Education India, 2006.

[3] Jacob Mathai, "History of Computer Security and Secrecy System". White Paper, pp. 3-8, April 2014.

[4] Rami Alnaqeib, Fahad H.Alshammari, M.A.Zaidan, A.A.Zaidan, B.B.Zaidan and Zubaidah M.Hazza, "An Overview of Extensible Markup Language Technology", *Journal of computing*,Vol2,Issue 6 ,pp.177-181,June 2010.

[5] Seifedine Kadry, "Document Security Using XML Technology", Proceedings of *International Conference on Electronic Engineering and Computer Science*, Publised by Elsevier B.V,pp. 1-6, 2013.

[6] Verma O.P., Agarwal R., Dafouti D. and Tyagi, S., "Performance Analysisof Data Encryption Algorithm", Proceedings of 3rd International Conference on Electronics Computer Technology, Vol. 5, pp. 399-403. April 2011.

[7] Gottesman, D., "Private key and public key quantum cryptography", Proceedings of IEEE International Conference on Quantum Electronics and Laser Science, pp. 309-315, May 2002.

[8] Quist-Aphetsi Kester, "Cryptosystem based on Vigenere cipher with varying key" *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, Vol 1,Issue 10,pp.108-113,December 2012.

[9] Elisa Bertino a,, Silvana Castano a, Elena Ferrari b and Marco Mesiti c , "Protection and administration of XML data sources", Published by Elsevier,pp 237-260 ,June 2002.

[10] El-Aziz, A.E.-A.A.A.and Kannan A., " A Cryptosystem for XML Document", *Computer Communication and Informatics (ICCCI)*, pp 1-3,Janunary 2012.

[11] NianLiu,Yajian Zhou ,Xinxin Niu and Yixian Yang, " A Novel Model for query over Encrypted XML daatabase" , Proceedings of IEEE International Conference on *Network Infrastructure and Digital Content*, pp 986-990, November 2009.

[12] Almarimi, A.and Alsahdi,U., "Developing Cryptosystem for XML Documents" ,Proceedings of 2nd International Conference on *Computer Technology and Development (ICCTD)*,pp-240-244, November 2010

[13] Jammalamadaka, R.C.and Mehrotra, S.,"Querying Encrypted XML Documents", Proceedings of 10th International conference on *Database Engineering and Applications Symposium*, pp 129-136,December 2006.

[14] Junqi Zhang ,Varadharajan, V.and Yi Mu, " Securing XML Document Sources and Their Distribution", Proceedings of 18th International Conference on *Advanced Information Networking and Applications*,Vol 1,pp-562-567,2004

[15] Nithin N and Anupkumar M Bongale, "XBMRSA:A New XML Encryption Algorithm.", Proceedings of Information and Communication Technologies (WICT), 2012 World Congress,pp 567-571,2012.

[16] [online] Caesar cipher -en.wikipedia.org/wiki/Caesar_cipher. 2003

[17] [online] Vigenere cipher-www.counton.org/explorer/codebreaking/vigenere-cipher.php. 2005